Faculty of Science

MASC06, Mathematical statistics: Data Analysis, Statistical Learning and Visualization with Project, 7.5 credits

Matematisk statistik: Dataanalys, statistisk inlärning och visualisering med projekt , 7,5 högskolepoäng

First Cycle / Grundnivå

Details of approval

The syllabus was approved by The Education Board of Faculty of Science on 2025-05-30. The syllabus comes into effect 2025-05-30 and is valid from the spring semester 2026.

General information

The course is an elective course for first-cycle studies for a degree of Bachelor of Science in mathematics but can also be read by students on other Bachelor programs in Scince. The course is also given as a stand alone course.

Language of instruction: English

Main field of study	*Specialisation*
Mathematics	G2F, First cycle, has at least 60 credits in first-cycle course/s as entry requirements
Mathematical Statistics	G2F, First cycle, has at least 60 credits in first-cycle course/s as entry requirements

Learning outcomes

The overarching goal of the course is to give an overview of basic data wrangling and visualisation, focusing on the ability to identify and illustrate important features of a data set.

Knowledge and understanding

On completion of the course, the student shall be able to:

- describe different ways of aggregating, summarising and visualising data,
- explain the principles of supervised learning,

- explain the importance of evaluating models based on their predictive ability.

Competence and skills

On completion of the course, the students shall be able to:

- wrangle, present and visualise data to highlight important features in a complex data material,
- use common methods for supervised learning,
- draw conclusions regarding a data material, based on results from classification and regression methods,
- use common method for evaluation of predictive ability on out-of-sample data,
- present the analysis and conclusions of a practical problem in a written report, by the help of literate programming.

Judgement and approach

On completion of the course, the students shall be able to:

- reflect over the limitations of the chosen model and method, as well as alternative solutions.
- reflect over the possible issues with fitting multiple models to the same data material.

Course content

The course treats:

- Basic methods for data handling and common visualisation methods for data
- Methods for unsupervised and supervised learning such as: clustering; hierarchical clustering; and regression and decision tree methods for classification and regression problems.
- Methods for model selection and validation such as: bootstrap, split of data into training and test, and cross-validation.

Course design

Teaching consists of lectures, computer exercises with written reports and a final project.

Participation in computer exercises is compulsory (see details under examination).

Assessment

The examination is done in form of written lab reports, peer review of reports, and projectreport with oral presentation, as well as attendance on at lest half of the scheduled labs, where of at least on scheduled lab per report.

Students who did not pass an assessment in the regular session will be offered another opportunity for assessment during the scheduled period for resits.

The examiner, in consultation with Disability Support Services, may deviate from the regular form of examination in order to provide a permanently disabled student with a form of examination equivalent to that of a student without a disability.

Grades

Grading scale includes the grades: Fail, Pass, Pass with distinction
The grade on the computer exercises with written report is Fail, Pass while the project is graded as Fail,Pass and Passed with distinction.

For a Pass grade on the whole course the student must have Pass grades on all written lab reports, peer review of reports, as well as attendance on at lest half of the scheduled labs, where of at least one scheduled lab per report.

The final grade is determined by the final project.

Entry requirements

For admission to the course, general entry requirements and 60 ECTS credits in Science, Technology, Engineering and/or Mathematics including knowledge equivalent to at least on of the flollowing courses:

MASA03 Mathematical statistics: Basic course, 15 ECTS,

MASB13 Mathematical Statistics for Physicists 7.5 ECTS,

MASB02 Mathematical Statistics for Chemists 7.5 ECTS.

English course 6/B.

Further information

The course is given by Centre for Mathematical Sciences, Lund University.

The course is read together FMSF90 Data Analysis: Statistical Learning and Visualization with Project which is a course given by Lund's engineering school LTH.

The examination of the course is scheduled according to LTH:s exam schedule.